MANUEL ARTIS ORTUÑO

Sobre la estimación de los modelos econométricos heteroscedásticos

(Método trietápico)

I. Planteamiento teórico

Consideremos el modelo econométrico heteroscedástico siguiente:

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + u_i \qquad [1]$$

en el que Y_i es la variable endógena, $X_{1i}, X_{2i}, \ldots, X_{ki}$ son las variables predeterminadas, $\beta_1, \beta_2, \ldots \beta_k$ son los parámetros y u_i es un término de perturbación que se distribuye normalmente con independencia de los valores de las variables predeterminadas.

A partir de un conjunto de N observaciones realizadas temporal o atemporalmente sobre las variables incluidas, y resumido en las matrices:

$$X = \left\| \begin{matrix} X'_1 \\ X'_2 \\ \vdots \\ X'_N \end{matrix} \right\| = \left\| \begin{matrix} X_{11}, & X_{21}, & \ldots, & X_{k1} \\ X_{12}, & X_{22}, & \ldots, & X_{k2} \\ \vdots & \vdots & & \vdots \\ X_{1N}, & X_{2N}, & \ldots, & X_{kN} \end{matrix} \right\| \qquad Y = \left\| \begin{matrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{matrix} \right\| \qquad [2]$$

en la primera de las cuales, cada fila X'_i indica el valor observado para el conjunto de las K variables predeterminadas en la i-ésima observación, y en la segunda, Y_i el valor de la variable endógena en ese mismo momento, es posible obtener estimaciones de los parámetros

$$\hat{\beta}' = \| \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k \|$$

si previamente se especifica un conjunto de hipótesis en relación con el comportamiento de las perturbaciones aleatorias u_i. Supongamos en ese sentido que se cumple que la variable tipificada $\{\sigma_i^{-1}u\}$ se distribuye siguiendo una distribución normal reducida, y que además no existe correlación serial entre las distribuciones correspondientes a los diferentes puntos de observación. Formalmente:

$$E\{u_i\} = \sigma_i E\{\sigma_i^{-1}u_i\} = 0 \qquad [3]$$

$$\text{Var}\,\{u_i\} = E\{u_i - 0\}^2 = \sigma_i^2 E\{\sigma_i^{-2}u_i^2\} = \sigma_i^2 \qquad [4]$$

y

$$\text{Covar}\,\{u_i u_{i'}\} = E\{\sigma_i^{-1}u_i\sigma_{i'}^{-1}u_{i'}\} = \sigma_i^{-1}\sigma_{i'}^{-1}\ E\{u_i u_{i'}\} = 0 \qquad [5]$$

Las dos últimas hipótesis pueden resumirse utilizando notación compacta en

$$E\{uu'\} = V \qquad [6]$$

donde $u = \|\, u_1, u_2, \ldots, u_N \,\|$ es el vector de perturbaciones correspondiente a los valores observados para las variables, y V es una matriz diagonal de la forma:

$$V = \begin{Vmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_N^2 \end{Vmatrix} \qquad [7]$$

El modelo anterior es heteroscedástico al no aceptarse la hipótesis suplementaria de que V sea una matriz escalar. Si fuese así el modelo sería homoscedástico, cumpliéndose como hipótesis alternativa a [6]:

$$E\{uu'\} = S = \sigma^2 I$$

que, en definitiva, se traduce en que la varianza de la perturbación sería en tal caso constante para cualquier observación, es decir,

$$\text{Var}\,\{u_i\} = E\{u_i^2\} = \sigma^2$$

Como señala Lozano (1965), la hipótesis de homoscedasticidad parece posible aceptarla para modelos que contengan variables agregadas observadas a lo largo del tiempo, toda vez que en tal caso los valores que toman las variables predeterminadas son de cuantía similar en todos los puntos de observación, y lo mismo ocurre con los valores de la variable endógena. Así, por ejemplo, en una función de consumo agregada, el consumo en períodos recientes no será sensiblemente distinto del nivel agregado de unos años atrás, y lo mismo ocurrirá con el nivel de renta. Ahora bien, cuando se manejan variables microeconómicas, la hipótesis de homoscedasticidad ya no resulta tan fácil de aceptar, pues los valores observados están afectados de sensibles diferencias en cuanto a magnitud. Son, entre otros, los casos de la función de gasto para una unidad familiar como ponen de manifiesto Prais y Houthakher (1955, pp. 55-58), y de inversión empresarial como muestran Meyer y Kuh (1957). Para éstos y otros similares resulta más apropiado el modelo heteroscedástico definido primeramente.

En un modelo del tipo que nos ocupa, la utilización de mínimos cuadrados ordinarios (MCO) para estimar los parámetros a través de la relación:

$$\hat{\beta} = (X'X)^{-1} X'Y \qquad [8]$$

da lugar a estimaciones que si bien son insesgadas resulta que no son eficientes según ha mostrado entre otros Johnston (1972, pp. 215-217) y el propio Lozano (1965), dependiendo la ineficiencia de la forma concreta como varíe σ_i^2 a lo largo del conjunto de observaciones. De ahí que, si como cabe esperar, se pretenden obtener estimaciones aceptables de los parámetros del modelo, sea imprescindible tomar en consideración de alguna manera la varianza de las perturbaciones. Una manera de lograrlo es mediante la utilización del método de los mínimos cuadrados generalizados (MCG), que da lugar al vector de estimaciones:

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}Y$$

y una matriz de varianzas y covarianzas

$$\text{Var }(\hat{\beta}) = (X'V^{-1}X)^{-1} \qquad [10]$$

de los que se obtiene como expresión detallada:

$$\hat{\beta} = \{\sum_{i=1}^{N} \sigma_i^{-2} X_i X'_i\}^{-1} \{\sum_{i=1}^{N} \sigma_i^{-2} X_i Y_i\} \qquad [9a]$$

$$\text{Var }(\hat{\beta}) = \{\sum_{i=1}^{N} \sigma_i^{-2} X_i X'_i\}^{-1} \qquad [10a]$$

cuando se tiene en cuenta la forma de las matrices [2] y [7] que intervienen.

II. Estimación de la matriz V

Los resultados tanto teóricos como empíricos obtenidos por los diversos autores —y en este último sentido es particularmente relevante la valoración que para una muestra de tamaño cinco hace Johnston de la pérdida de eficiencia debida a la utilización indebida de MCO, la cual cifra entre 30 y un 50 por ciento— sugieren que la heteroscedasticidad es un problema realmente serio que no puede abordarse sin la utilización de MCG. Ahora bien, ello no es tan sencillo como podría pensarse, ya que tal utilización tropieza con la dificultad de que en las fórmulas [9] o [9*a*] aparecen las σ_i^2 que son desconocidas, dificultad agravada por el hecho de que normalmente resulta difícil averiguar de qué modo se comportan, e incluso saber cuáles son las causas que motivan su variación. En tales circunstancias la única solución que queda es estimarlas a partir de un modelo que explique su comportamiento. Modelo que se establece a base de postular que σ_i^2 puede explicarse mediante una relación funcional del tipo:

$$\sigma_i^2 = f\{Z_{1i},\ Z_{2i},\ \ldots,\ Z_{pi};\quad \alpha_1,\ \alpha_2,\ \ldots,\ \alpha_p\} \qquad [11]$$

en la que Z_{1i}, Z_{2i}, ..., Z_{pi}, son P variables predeterminadas, que pueden coincidir totalmente o en parte —sin que ello sea imprescindible—, con las X_{Ki} de [1]. Las α_1, α_2, α_3, ..., α_p son parámetros a estimar y la característica funcional f varía según los diferentes autores que han abordado la formulación de este modelo auxiliar.

Antes de pasar a describir las tres líneas más destacadas de actuación en orden a estimar [1] mediante la especificación de [11] es preciso hacer constar que al proceder tal como indicamos se lograrán diversas estimaciones alternativas de V que podrán utilizarse por analogía con [9] y [10] para estimar β mediante

$$b = (X'\hat{V}^{-1}X)^{-1}\,(X'\hat{V}^{-1}Y) \qquad [12]$$

cuya matriz de varianzas y covarianzas será:

$$\text{Var}\ (b) = (X'\hat{V}^{-1}X)^{-1} \qquad [13]$$

A pesar de la analogía formal, ni las estimaciones ni su eficiencia coincidirán con las de $\hat{\beta}$, de lo que se sigue que será indispensable analizar para cada método que se proponga las propiedades que se obtienen para b en relación con las de $\hat{\beta}$.

A) *Estimaciones bietápicas*

Dentro de una primera línea destacan las aportaciones de Geary (1966), Park (1966) y Glejser (1969), los cuales proponen estimar β mediante las relaciones [12] y [13] en las que $\hat{V}$ está formada por los valores σ_i^2 que se estiman al efectuar la regresión deducida de una especificación lineal del modelo [11], en la que como valores del regresando se toman los errores de estimación e_i^2 obtenidas al aplicar a [1] MCO. La forma más habitual que adopta [11] es la correspondiente a alguno de los tres tipos siguientes:

$$
\begin{array}{ll}
a) & \sigma_i = \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \ldots + \alpha_p Z_{pi} \\
b) & \sigma_i^2 = \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \ldots + \alpha_p Z_{pi} \\
c) & \log \sigma_i^2 = \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \ldots + \alpha_p Z_{pi}
\end{array}
\qquad [14]
$$

Para fijar ideas, y sin que ello suponga pérdida de generalidad, en lo que sigue nos centraremos en la especificación *b*) de las anteriores. Resultará:

$$e_i^2 = \alpha_1 Z_{1i} + \alpha_2 Z_{2i} + \ldots + \alpha_p Z_{pi} + v_i \qquad [15]$$

en donde v_i son las variables aleatorias:

$$v_i = e_i^2 - E\{u_i^2\} \qquad [16]$$

y e_i^2 los cuadrados de los elementos del vector

$$e' = \| e_1, e_2, \ldots, e_N \| = (Y - X\beta)'$$

deducido de [8].

La esencia del método que consideramos consiste en obtener una estimación

$$\hat{\alpha}' = \| \hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_p \|$$

que permita escribir:

$$\sigma_i^2 = \hat{\alpha}_1 Z_{1i} + \hat{\alpha}_2 Z_{2i} + \ldots + \hat{\alpha}_p Z_{pi} = Z_i' \hat{\alpha} \qquad [17]$$

en donde Z_i es el vector de los valores observados en el punto i para el conjunto de las P variables explicativas en [11], completándose la estimación de $\boldsymbol{\beta}$ con la sustitución de [17] en [12] que adopta la expresión particular

$$b_1 = \{ \sum_{i=1}^{N} (Z_i' \hat{\alpha})^{-1} X_i X_i' \}^{-1} \{ \sum_{i=1}^{N} (Z_i' \hat{\alpha})^{-1} X_i Y_i \} \qquad [18]$$

La crítica que puede formularse al anterior proceso, según indican Goldfeld y Quandt (1972, p. 93) para el modelo de Glejser —especificación *a*) de las anteriores— pero que es de fácil generalización al caso que nos ocupa, hace referencia a la validez de los parámetros $\hat{\alpha}$. Desafortunadamente el término de perturbación [16] posee un elevado número de propiedades indeseables, a saber, es de esperanza no nula, está autocorrelacionado y lo que ya parece que colma sus inconvenientes es que también es heteroscedástico. De ahí que cuando se proceda a obtener como estimación de α la correspondiente al método de los mínimos cuadrados ordinarios, es decir:

$$\hat{a} = \left\{\sum_{i=1}^{N} Z_i Z'_i\right\}^{-1} \left\{\sum_{i=1}^{N} Z_i e_i^2\right\} \qquad [19]$$

su significación sea realmente escasa. A pesar de lo cual vale la pena hacer constar que toda vez que $\{e_i^2 - u_i^2\}$ converge en probabilidad hacia cero, la perturbación v_i convergerá a su vez en distribución hacia

$$v_i^* = u_i^2 - E\{u_i^2\} \qquad [20]$$

que, como fácilmente se comprueba, se distribuye con media nula y sin autocorrelación serial, motivo por el cual $\hat{a}$ será, a pesar de todo, una estimación consistente y asintóticamente insesgada de α, con matriz de varianzas y covarianzas

$$\text{Var}\ (\hat{a}) = \left\{\sum_{i=1}^{N} Z_i Z'_i\right\}^{-1} \qquad [21]$$

La causa de tal consistencia es la que, según demuestra Harvey (1974), determina también que la estimación b_1 se distribuya asintóticamente con análogas propiedades que la obtenida por mínimos cuadrados generalizados, y, de ahí, que en principio pueda pensarse en utilizarla para estimar β ya que, para tamaños muestrales elevados, los resultados serán aceptables.

Ahora bien, como cualquier persona que se dedique a la Econometría sabe, en la práctica y con mayor frecuencia de lo deseable, es corriente disponer solamente de un conjunto escaso de observaciones sobre las variables del modelo. De ser así, el sesgo de $\hat{a}$ afectará las estimaciones de b_1. Para tales casos parece recomendable utilizar un procedimiento iterativo que aparece citado constantemente por los autores que analizan este tema y que consiste en utilizar los errores de estimación $\hat{a}_1$ de α, que sustituida en [18] determine a su vez una nueva b_2 de β, y así sucesivamente hasta que la diferencia entre dos estimaciones consecutivas b_N y b_{N+1} sea prácticamente nula, con lo cual y bajo determinadas condiciones bastante generales de convergencia, pueden irse mejorando las sucesivas estimaciones.

Hay que hacer notar, por último, que la varianza asintótica de la perturbación v_i será

$$\text{Var}\ \{v_i^n\} = E\{v_i^x\}^2 = E\{u_i^2 - \sigma_i^2\} = E\{u_i^4\} - \sigma_i^4 = \sigma_i^4\{\mu_4 - 1\} = 2\sigma_i^4 \qquad [22]$$

donde μ_4 es el momento de orden cuatro de una distribución normal reducida que vale siempre 3. Al no ser [22] constante, muestra cómo la heteroscedas-

5

ticidad sigue presente en el modelo [15] por más que se aumente el número de observaciones sobre las variables, lo cual lleva consigo el que las estimaciones de α obtenidas por mínimos cuadrados ordinarios sean ineficientes, incluso asintóticamente.

B) *Estimaciones maximoverosímiles*

Otra línea de aproximación al problema de estimar los modelos heteroscedásticos es la inicialmente propuesta por Theil (1951). Consiste ésta en estimar el modelo simultáneamente en lugar de hacerlo en dos etapas (primero $\hat{\alpha}$ y después $\hat{\hat{\beta}}$) como antes, utilizando para ello el método de la máxima verosimilitud (MV). La función de verosimilitud que habrá que usar para ello será

$$L\alpha\{\sqrt{2n}\}^{-N/2} \mid V \mid^{-1/2} e^{-1/2(Y-X\beta)'V^{-1}(Y-X\beta)} \qquad [23]$$

en la que si admitimos para σ_i^2 la misma especificación que en el apartado anterior, es decir,

$$\sigma_i^2 = Z_i'\alpha$$

V tomará la forma:

$$V = \begin{Vmatrix} Z_1'\alpha & & & & \\ & Z_2'\alpha & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & Z_N'\alpha \end{Vmatrix} \qquad [24]$$

de la que se obtiene la siguiente especificación detallada para [23]:

$$L\alpha \prod_{i=1}^{N} \{2\pi \{2\pi\}^{-1/2} (Z_i'\alpha)^{-1/2} e^{-1/2[(Y_i - X_i')^2/Z_i'\alpha]} \qquad [23a]$$

a partir de la cual son deducibles las estimaciones $\hat{\alpha}$ y $\hat{\beta}$.

La mayor dificultad con la que tropieza la aplicación del presente método es su falta de operatividad, debida a que los algoritmos que se deducen de [23] como estimaciones de α y β son muy difíciles de computar. Para paliar en parte este problema, Rutemiller y Bowers (1968) han propuesto un método iterativo, sobre la idea del «method of scoring» de Rao (1965, pp. 302-309), mediante el cual es posible aproximar las estimaciones de β, y en la misma línea de investigación Levenbach (1973) propone la estimación de los parámetros mediante las funciones marginales de verosimilitud. Hay que resaltar, de todos modos, que ambas soluciones siguen siendo complicadas de manejar y, por lo demás, no están suficientemente claras las condiciones de convergencia sobre las que están basadas.

C) *Método de Theil-Prais*

El método que pasamos a esbozar —que en realidad es un caso particular del método bietápico antes considerado— se halla inmejorablemente expuesto en Theil (1971, pp. 244-248 y 403-405) y está basado en los trabajos de Prais (1953) y Prais y Houthakker (1955). El que lo hayamos considerado por separado se debe fundamentalmente a dos razones: a que históricamente ha sido el primero en utilizarse en la estimación de modelos heteroscedásticos, lo cual parece justificar un posible trato diferenciado, por una parte, y por otra a que más adelante haremos uso de él en este trabajo, por lo que vale la pena conocer exhaustivamente su funcionamiento.

La idea básica sobre la que se sustenta consiste en admitir que la varianza de las perturbaciones es proporcional al cuadrado de la esperanza matemática de la variable endógena dentro del modelo que se está manejando. Formalmente la hipótesis que debe satisfacerse es:

$$\sigma_i^2 = K[E(Y_i)]^2 \qquad [25]$$

Bajo tal premisa, si escribimos de nuevo el modelo inicial, ahora bajo la forma:

$$Y_i = X_i'\beta + u_i \qquad [26]$$

se cumplirá

$$E\{Y_i\} = X_i'\beta \qquad [27]$$

de donde, y según [25]:

$$\sigma_i^2 = K\{X_i'\beta\}^2 \qquad [28]$$

que permite concluir que la estimación por mínimos cuadrados generalizados del modelo sería, en este caso:

$$\hat{\beta} = \{\sum_{i=1}^{N} K^{-1}(X_i'\beta)^{-2}X_iX_i'\}^{-1} \quad \{\sum_{i=1}^{N} K^{-1}(X_i'\beta)^{-2}X_iY_i\} \qquad [29]$$

o lo que es lo mismo por simplificación:

$$\hat{\beta} = \{\sum_{i=1}^{N} (X_i'\beta)^{-2}X_iX_i'\}^{-1} \quad \{\sum_{i=1}^{N} (X_i'\beta)^{-2}X_iY_i\} \qquad [29a]$$

que es independiente de la constante de proporcionalidad. Análogamente, su matriz de varianzas y covarianzas será:

$$\mathrm{Var}\,\{\hat{\beta}\} = K\ \{\sum_{i=1}^{N} (X_i'\beta)^{-2} X_i X_i'\}^{-1} \qquad [30]$$

Como fácilmente puede constatarse, la estimación anterior presupone el conocimiento del vector β que es precisamente el que queremos estimar. Por ello, un requisito previo para su aplicación es haber estimado éste por mínimos cuadrados ordinarios, para obtener entonces una fórmula del tipo:

$$b_1 = \ \{\sum_{i=1}^{N} (X_i'\hat{\beta})^{-2} X_i X_i'\}^{-1}\ \{\sum_{i=1}^{N} (X_i'\hat{\beta})^{-2} X_i Y_i\} \qquad [31]$$

con

$$\mathrm{Var}\,\{b_1\} = K\ \ \{\sum_{i=1}^{N} (X_i'\hat{\beta})^{-2} X_i X_i'\}^{-1} \qquad [32]$$

que puede servir de base para la aplicación reiterada del método como ya se apuntó en A).

Es preciso resaltar que, dada la forma como se obtienen los resultados, éstos deben coincidir, según [28] con los que aparecerían al aplicar el método bietápico cuando se tomase como especificación de [11] la *a*) de [14], en la que las variables explicativas Z_{si} fuesen precisamente las X_{ri} del modelo [1], y en el caso, además, de que el vector de parámetros mostrase proporcionalidad con β. Esta similitud de resultados es la razón de que [31] goce análogamente de las mismas propiedades asintóticas que las obtenidas por mínimos cuadrados generalizados.

En relación con este método existe una variante indicada por Lozano (1965) mediante la cual se amplían enormemente sus posibilidades de aplicación. Dicha variante consiste en aceptar como hipótesis de comportamiento para la perturbación la más general

$$\sigma_i^2 = \mathrm{K}[E(Y_i)]^4 \qquad [33]$$

obteniéndose en dicho caso las siguientes estimaciones:

$$b_1 = \ \{\sum_{i=1}^{N} (X_i'\hat{\beta})^{-4} X_i X_i'\}^{-1}\ \{\sum_{i=1}^{N} (X_i'\hat{\beta})^{-4} X_i Y_i\} \qquad [31a]$$

con

$$\mathrm{Var}\,(b_1) = K\ \{\sum_{i=1}^{N} (X_i'\hat{\beta})^{-4} X_i X_i'\}^{-1} \qquad [32a]$$

cuyas propiedades siguen siendo las mismas.

Queda, por último, hacer mención del trabajo de Ameniya (1973) en el cual se realiza un esfuerzo de sistematización de los resultados alcanzados por este método que es realmente importante.

D) *Resultados experimentales*

Ya se ha indicado que en la práctica no es raro disponer de información escasa sobre el comportamiento de las variables presentes en el modelo. En tales casos, el que las estimaciones que se pueden obtener aplicando cualquiera de los métodos detallados sean asintóticamente eficientes es un triste consuelo. Por ello, es imprescindible analizar el comportamiento que muestran todos ellos para tamaños reducidos, y esto es sabido que sólo se logra simulando experimentalmente su actuación para tales tamaños. Esto es precisamente lo que hacen Goldfeld y Quandt (1972, pp. 102-120) los cuales para cada una de las nueve estructuras distintas que seleccionan en relación con el modelo

$$Y_i = \alpha + \beta X_i + u_i \qquad \sigma_i^2 = a + bX_i + CX_i^2 \qquad [34]$$

obtienen sucesivamente las estimaciones bietápicas y maximoverosímiles que se detallaron en el apartado *A*) la primera y en el *B*) la segunda. Utilizando tamaños muestrales que varían de treinta a noventa en todos los casos, dichos autores realizan un número de estimaciones que en unos casos es cincuenta y en otros cien, obteniendo la eficiencia comparativa de los métodos seleccionados, para cada una de las estructuras y tamaños muestrales, medida a través del error cuadrático medio respecto a los parámetros verdaderos. Pues bien, la conclusión que obtienen de la experimentación anterior es que con mayor o menor intensidad ambos estimadores son ineficientes, lo cual ya era presumible, pero que comparativamente las estimaciones maximoverosímiles son significativamente más eficientes que las bietápicas.

Nos encontramos, según lo anterior, ante la siguiente situación. Por una parte, las estimaciones maximoverosímiles son mejores, pues dan estimaciones relativamente más eficientes y, sin embargo, como contrapartida son más penosas de obtener y en algunos casos prácticamente imposibles. Por tal motivo, lo único que está claro es que la elección es una cuestión delicada y desde luego susceptible de producir controversias.

III. ESTIMACIONES TRIETÁPICAS

El objetivo último del presente trabajo estriba en fundamentar un método de estimación de β, alternativo de los mencionados hasta ahora, que sea tan sencillo de aplicar como lo es el bietápico y que, sin embargo, mejore dentro de lo posible las estimaciones que éste proporciona. A tal fin se propone el siguiente:

$$\beta^* = \{\sum_{i=1}^{N} (Z'_i\alpha)^{-1} X_i X'_i\}^{-1} \{\sum_{i=1}^{N} (Z'_i\hat{\alpha})^{-1} X_i Y_i\} \qquad [35]$$

cuya diferencia esencial con [18] está en las $\hat{\alpha}$ utilizadas, para las cuales se exige en este caso que sean asintóticamente eficientes, propiedad que no poseen las $\hat{a}$ al ser estimaciones por mínimos cuadrados ordinarios del modelo heteroscedástico:

$$e_i^2 = \sigma_i^2 + v_i \qquad \text{con} \qquad \sigma_i^2 = Z'_i\alpha \qquad [36]$$

que, asintóticamente, recordemos adopta la forma

$$u_i^2 = \sigma_i^2 + v_i^* \qquad [37]$$

en la que

$$E\{v_i^*\} = 0 \qquad \text{Var}\ \{v_i^*\} = 2\sigma_i^4 \qquad [38]$$

La idea de seleccionar $\hat{\alpha}$ cumpliendo el requisito de ser asintóticamente eficiente más que un presupuesto para la estimación de β aparece como una consecuencia del rigor con que debe obtenerse también la estimación del modelo en que se definen las α dada su heteroscedasticidad. Por ello deberá ser tratado mediante cualquiera de los métodos desarrollados a lo largo de las páginas anteriores. Una posible crítica de la forma de actuación apuntada es la de denunciar que estamos en presencia de un proceso indefinido. En efecto, como para obtener $\hat{\alpha}$ es necesario conocer de algún modo el comportamiento de la varianza de v_i^*, parece que el siguiente paso será especificar previamente un nuevo modelo que la explique, el cual será también heteroscedástico, por lo que requerirá a su vez de otro, y así sucesivamente sin que jamás se llegue a una situación definitiva.

Por fortuna, lo anterior no es exactamente cierto y la verdad es que no hace falta ningún nuevo modelo, ya que al escribir la relación [38] bajo la nueva forma:

$$\text{Var}\ \{v_i^*\} = 2\ [E(u_i^2)]^2$$

se comprueba cómo el modelo [36] satisface *siempre* (y esto es lo verdaderamente importante) la hipótesis previa de aplicación del método de Theil, que recordemos exige que la varianza del término de perturbación sea proporcional al cuadrado de la esperanza matemática de la variable endógena. Por ello resulta sencillo obtener $\hat{\alpha}$, que adopta la forma

$$\hat{\alpha} = \{\sum_{i=1} (Z'_i\hat{a})^{-2} Z_i Z'_i\}^{-1} \{\sum_{i=1} (Z'_i\hat{a})^{-2} Z_i e_i^2\} \qquad [39]$$

donde e_i son los errores de estimación de β por mínimos cuadrados ordinarios, y α la de α por el mismo método.

El nombre de trietápicos proviene de que la estimación de β se obtiene en tres etapas, a saber, primero estimemos $\hat{\alpha}$, luego $\hat{\hat{\alpha}}$ por [39], y por fin β^* por [35], y la idea de su estructura está inspirada en Theil (1971, pp. 623-625) que desarrolla un estimador similar en el contexto de los modelos con parámetros estocásticos.

IV. Elección entre estimadores

El estimador β^* que se ha propuesto en el punto anterior se basa en la aplicación del método bietápico para unas estimaciones de α que sean asintóticamente eficientes. Ocurre, por tanto, que la ineficiencia de las $\hat{\hat{\alpha}}$ utilizadas desaparece para tamaños muestrales grandes, pero para los pequeños sigue presente, y si bien cabe pensar que en tal caso aún seguirán siendo mejores las estimaciones de β obtenidas por β^* que las obtenidas por cualquier otro método, ello no puede asegurarse si no es a la luz de alguna evidencia en dicho sentido. Como tal, proponemos la comparación de las estimaciones por mínimos cuadrados ordinarios [8], mínimos cuadrados bietápicos [18], y mínimos cuadrados trietápicos [35], del modelo con una sola variable exógena:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad [40]$$

obtenidas a partir de los valores X_i e Y_i ($i = 1, 2, \ldots, 78$) que se incluyen al final del trabajo. Dichos valores han sido generados aleatoriamente mediante el siguiente mecanismo por medio del cual se garantiza la heteroscedasticidad en el modelo: En primer lugar se generan los 78 valores para la variable exógena, a partir del supuesto de que se distribuyen uniformemente. A continuación se obtienen los valores de la perturbación para cada uno de los valores observados para la variable anterior, mediante la relación

$$u_i = \{20 + X_i + 0{,}25\, X_i^2\}^{1/2}\, \zeta_i \qquad [41]$$

en la que ζ_i es una variable aleatoria que sigue una distribución normal reducida y ya por fin se obtienen los valores de la variable endógena, a través de [40], bajo la hipótesis de que la estructura que los genera es aquella cuyos parámetros son

$$\beta_1 = \beta_2 = 2 \qquad [42]$$

Nótese que [41], junto con la forma de la distribución de ζ_i implican:

$$\sigma_i^2 = \{20 + X_i + 0{,}25\, X_i^2\} \qquad [43]$$

indicativo de cuáles han sido los parámetros α que junto con [42] completan la estructura seleccionada, e informa sobre el tipo particular de relación funcional escogido para explicar el comportamiento de la varianza de las perturbaciones.

La estimación por mínimos cuadrados ordinarios obtenida para el modelo es:

$$\text{(MCO)} \qquad \hat{Y}_i = \underset{(0{,}9937)}{1{,}79575945} + \underset{(1{,}6186)}{2{,}22483052}\, X_i$$

la estimación por mínimos cuadrados bietápicos alcanzada mediante dos iteraciones cuyos valores fueron, para b_{11} 1,85025107 y 1,85274292, y para b_{12} 2,20234779 y 2,20093415 resulta ser:

$$\text{(MC2E)} \qquad \hat{Y}_i = \underset{(0{,}9551)}{1{,}85285071} + \underset{(1{,}5426)}{2{,}20076895}\, X_i$$

y la estimación por mínimos cuadrados trietápicos

$$\text{(MC3E)} \qquad \hat{Y}_i = \underset{(0{,}9511)}{1{,}86114305} + \underset{(1{,}5349)}{2{,}19691293}\, X_i$$

resultando este último susceptible de ser mejorado, pues ha sido obtenido sin ninguna iteración posterior.

A pesar de la validez relativa de las conclusiones que se pueden derivar de la observación conjunta de los tres resultados, que deberían verse corroboradas por la aplicación reiterada del mecanismo descrito, para diversas estructuras, e incluso para diversos modelos especificados alternativamente, puede afirmarse que tales resultados no están en contradicción con los planteamientos teóricos previos, y en cualquier caso, dejan una puerta abierta a una posible vía de investigación futura.

Facultad de Ciencias Económicas
Universidad de Barcelona

BIBLIOGRAFÍA

1. AMENIYA, T.: «Regression Analysis when the Variance of the Dependent Variable is Proportional to the Square of its Expectation», *J.A.S.A.*, 68, 1973, pp. 928-934.
2. GEARY, R. C.: «A Note on Residual Heterovariance and Estimation Efficiency in Regression», *American Statistician*, 20, 1966, pp. 30-31.
3. GLEJSER, H.: «A new Test for Heteroscedasticity», *J.A.S.A.*, 64, 1969, pp. 316-323.
4. GOLDFELD, A. M., y QUANDT, R.E.: *Nonlinear Methods in Econometrics*, North Holland Publishing Company, Amsterdam, 1972.

5. HARVEY, A. C.: «The Estimation of Parameters in a Heteroscedastic Regression Model», presentado al Congreso de Econometría Europeo, Grenoble, 1974.
6. JOHNSTON, J.: *Econometric Methods,* 2.ª ed., McGraw-Hill, Nueva York, 1972.
7. LEVENBACH, H.: «The Estimation of Heteroscedasticity from a Marginal Likelihood Function», *J.A.S.A.,* 68, 1973, pp. 436-439.
8. LOZANO, V.: «Modelos de Regresión Lineal con Perturbaciones Heteroscedásticas», *Estadística Española,* 1965, pp. 57-73.
9. MEYER, J. R., y KUH, E.: *The Investment Decision: an Empirical Study,* Harvard University Press, Cambridge, Mass., 1957.
10. PARK, R. E.: «Estimation with Heteroscedastic Error Terms», *Econometrica,* 34, 1966, p. 888.
11. PRAIS, S. J.: «A Note on Heteroscedastic Errors in Regression Analysis», *Review of the International Statistical Institute,* 21, 1953, pp. 28-29.
12. PRAIS, S. J., y HOUTHAKKER, H. S.: *The analysis of Family Budgets,* Cambridge University Press, 1955.
13. RAO, C. R.: *Linear Statistical Inference and Its Applications in Biometric Research,* John Wiley and Sons, Inc., Nueva York, 1965.
14. RUTEMILLER, H. C., y BOWERS, D. A.: «Estimation in a Heteroscedastic Regression Model», *J.A.S.A.,* 63, 1968, pp. 552-557.
15. THEIL, H.: «Estimates and Their Sampling Variance of Parameters of Certain Heteroscedastic Distributions», *Review of the International Statistical Institute,* 19, 1951, pp. 141-147.
16. THEIL, H.: *Principles of Econometrics,* John Wiley Sons, Inc., Nueva York, 1971

Valores utilizados para las estimaciones del modelo

Exógena (X)		Endógena (Y)	
0,38532257	0,34459741	3,15251399	— 6,83492610
0,57296789	0,91069057	3,85531979	6,58543899
0,82953753	0,33830389	2,13972864	11,74614625
0,71215981	0,78440953	9,79357339	4,84267633
0,27459537	0,08530381	9,80514499	0,36957283
0,50217349	0,30268337	— 6,47558389	8,23200978
0,22876873	0,45694949	5,65408014	1,09702937
0,53376221	0,73052073	— 0,67144234	5,14840054
0,95958017	0,93666621	— 2,89378440	9,19716166
0,10920309	0,11918817	6,11023237	— 5,46471592
0,23644793	0,87101909	— 1,69130562	— 2,13316907
0,23786061	0,24027993	4,21701906	— 2,43420688
0,06075697	0,02092461	2,68046594	7,39162702
0,49101669	0,93268497	0,85578521	1,58688684
0,95849513	0,98147269	8,48699613	2,36012468
0,26029501	0,82760713	5,39040214	0,88682557
0,69780577	0,18991901	1,99673523	5,41377744
0,98297429	0,89485377	7,07847175	1,54506997
0,45763033	0,58767029	8,86361912	— 0,41055948
0,91050541	0,28322233	7,36408721	10,00203466
0,65760657	0,80908941	— 2,73184224	2,42720344
0,09483589	0,10457457	6,06047822	0,77268280
0,82537353	0,22337189	4,91086968	— 2,98410214
0,91553181	0,54664553	2,72755414	7,43931680
0,28223937	0,90147581	1,95567874	1,63011626
0,88676149	0,69592737	— 2,52150837	2,03043010
0,13404473	0,93273749	6,54545104	3,22532376
0,72401421	0,45819673	— 4,60358917	2,68959286
0,87698417	0,05171821	3,27730609	— 0,30126930
0,38531109	0,44619217	3,77708341	— 2,05098550
0,46876393	0,08632709	0,60662592	6,68510993
0,97019261	0,59099593	4,79919120	5,93651259
0,43832097	0,47005661	8,88306201	6,42571747
0,63944469	0,70384897	3,93685274	1,22743679
0,23945113	0,83710069	2,35870711	9,34976424
0,49790701	0,36296313	5,40455761	1,84744654
0,50192977	0,61633101	— 5,69203050	— 2,17986782
0,01652229	0,43657777	0,49134484	— 0,68665998
0,97282633	0,81641829	6,65485848	8,03570525